# Machine Learning & Misinformation:

## How algorithms can help us identify “alternative facts”

bookface

Brian Mendes, Johann Baydeer, Elliot Richardson

# Why are we here today?

**Problem**

Determining truth of an article manually cannot keep up with the rate at which news articles are produced, shared, and cited.

**Proposal**

Utilize machine learning to identify patterns in fake news and apply a filtering algorithm to articles shared on Bookface.

**Recommendations**

Implement our model to assign credibility scores and optimize credibility thresholds to automatically label clearly real or fake news and minimize manual fact-checking.

bookface

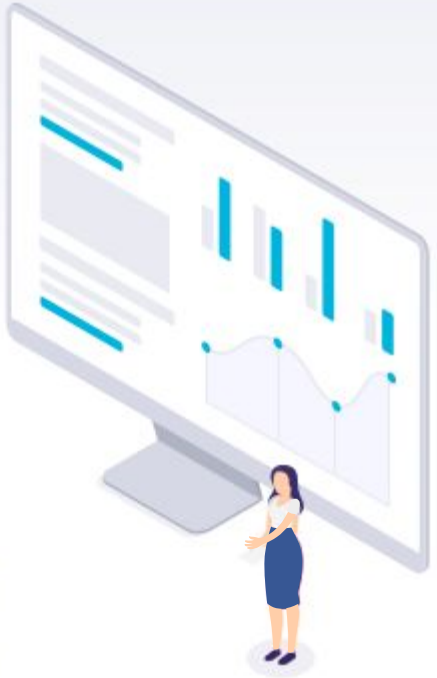
# Today, we'll go over:

**Process:**

- Character analysis
- Syntactical breakdown
- Sentiment analysis
- Significant words

**Results:**

- Exploratory findings
- Uninformative features
- Modeling results
- Recommendations



bookface

# Our process

- **Character analysis**
  - Sentence and word length
  - Punctuation
- **Syntactical breakdown**
  - Part-of-speech tagging
- **Sentiment analysis**
  - Positive, negative, neutral, & compound
- **Significant words**

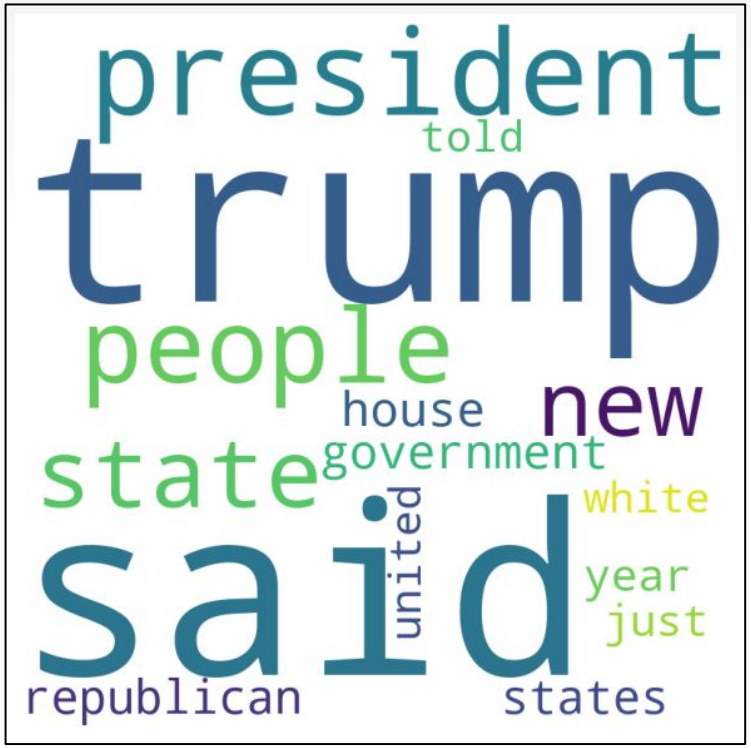
bookface

# Exploratory findings

- **Average length:** Fake articles are longer
  - 442 words for fake vs. 393 for real
- **Fake news articles:**
  - Negativity, interjections & symbols
- **Real news articles:**
  - Less negativity, more numbers & nouns
- **Time series** & **clustering** don't help
- Very **similar words** used
  - "Trump" in fake news, "President" in real news

bookface

# Exploratory findings



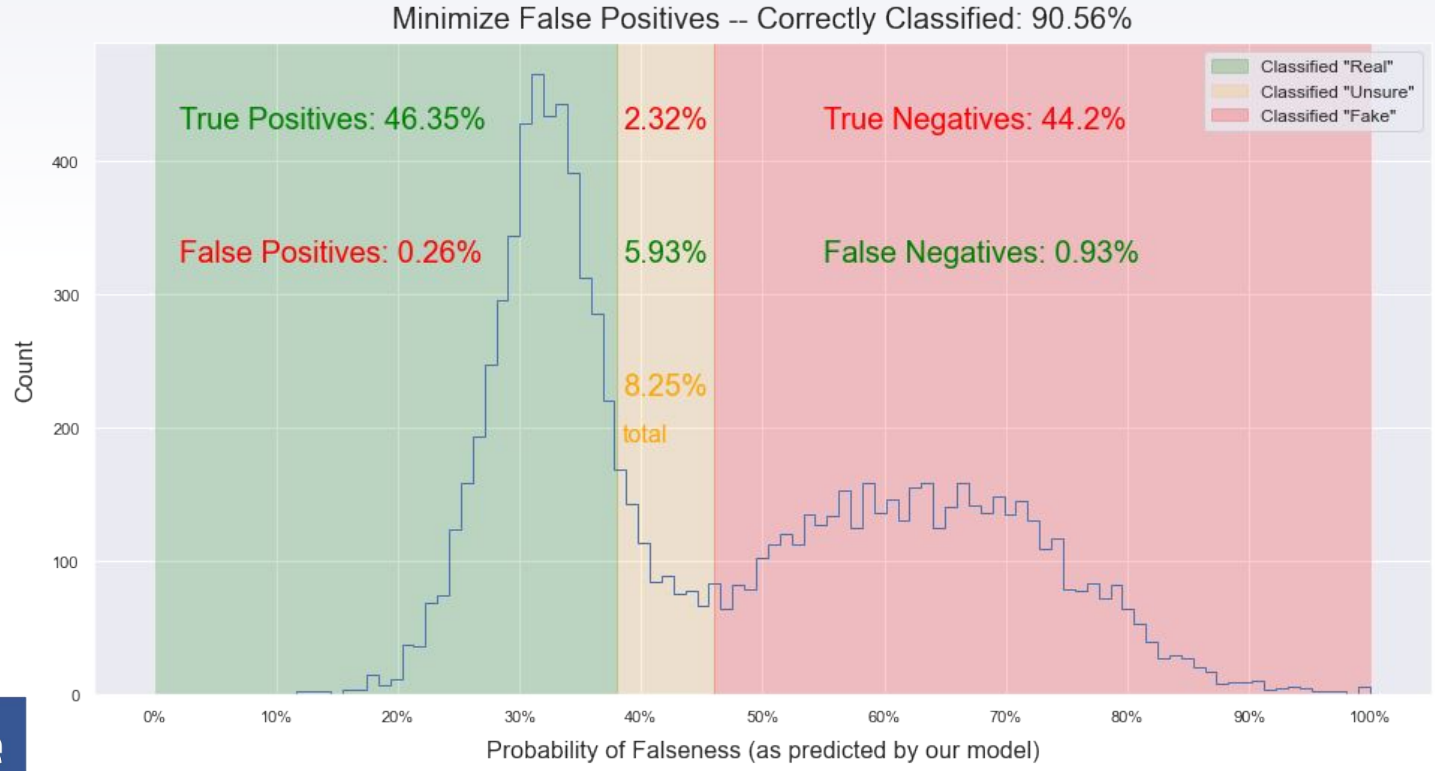

bookface

# Modeling results

- **Logistic regression model:**
  - Regularization & scaled data
  - Cross-validated score on test data: 85.9
  - 11% false negatives, 1% false positives
- **Recommendations for use:**
  - Optimize credibility thresholds

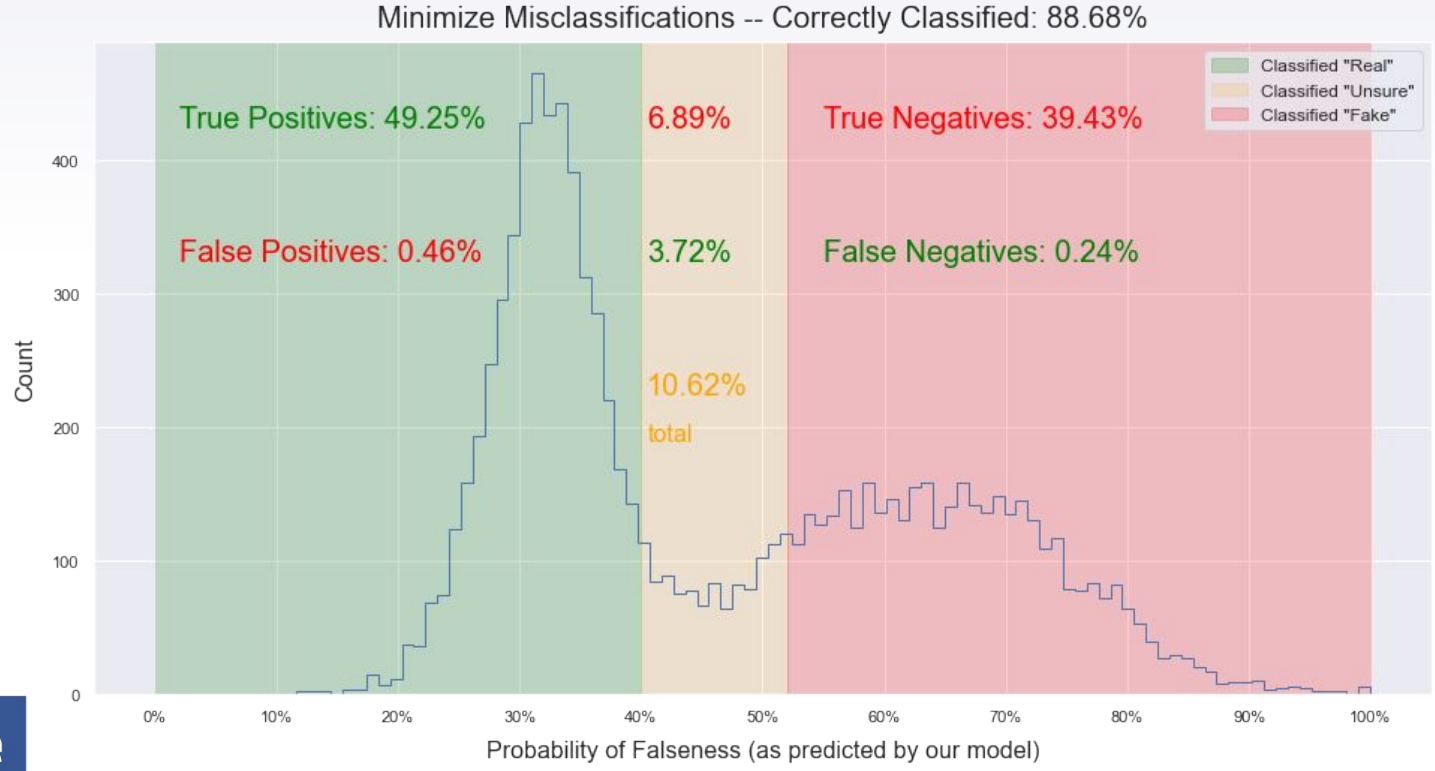
bookface

# Modeling results



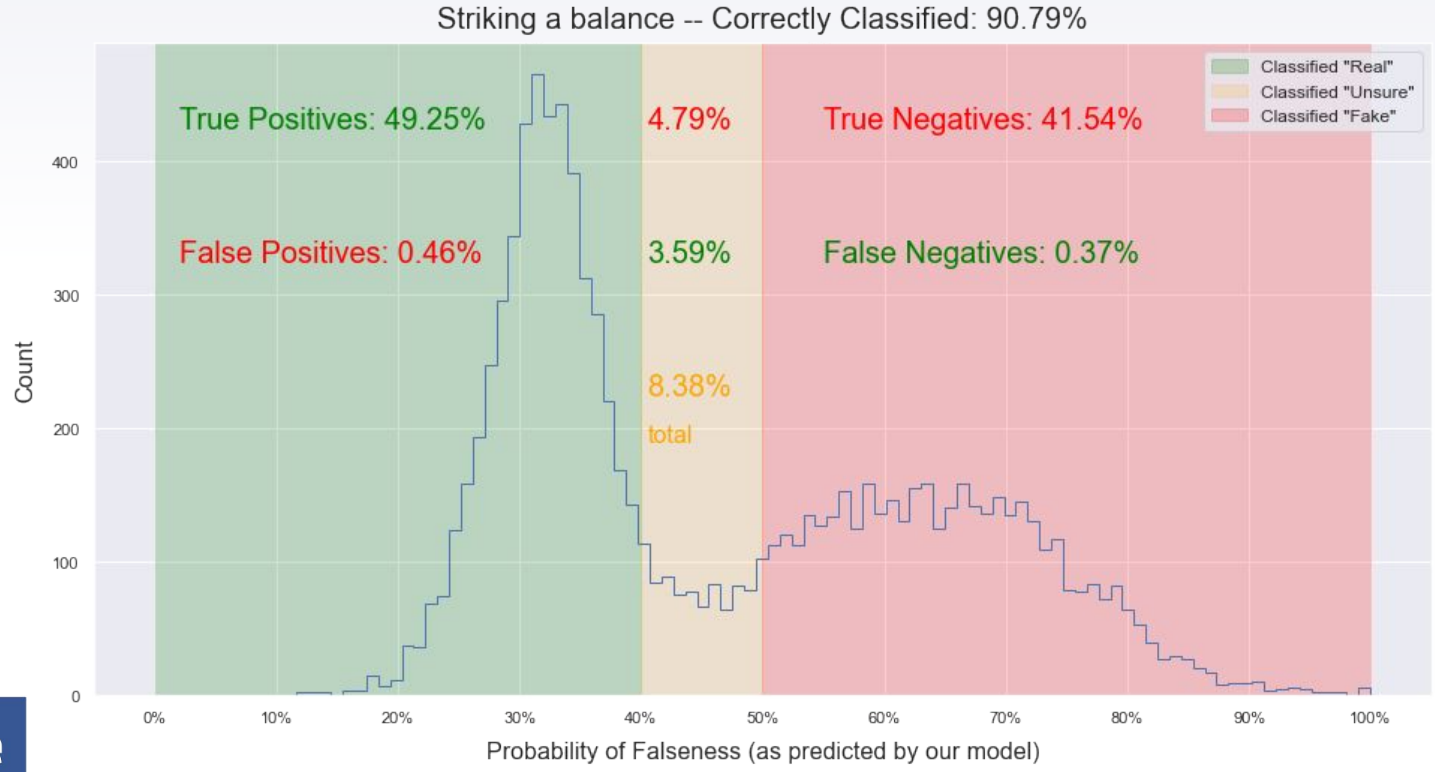

bookface

# Modeling results




bookface

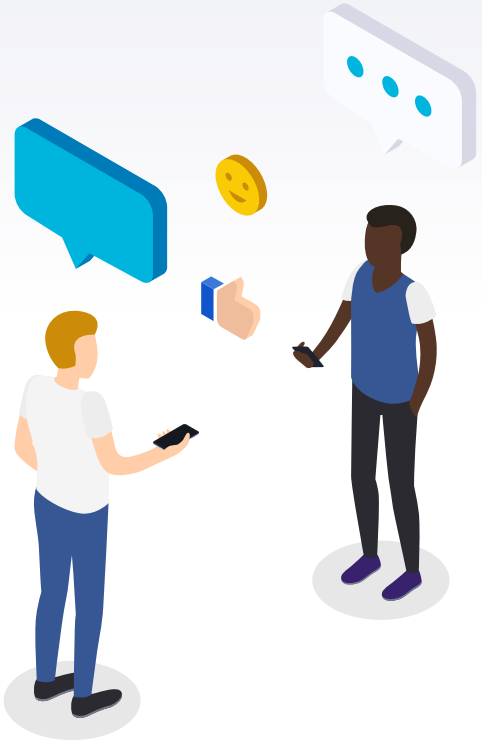
# Modeling results




bookface

# We can clean up Bookface together!

## Any questions?



bookface